Amplify Science is associated with better grade 5 science outcomes in Washington

By Leah Walker, Ph.D.
Executive Director Assessment and Reporting at Amplify Education

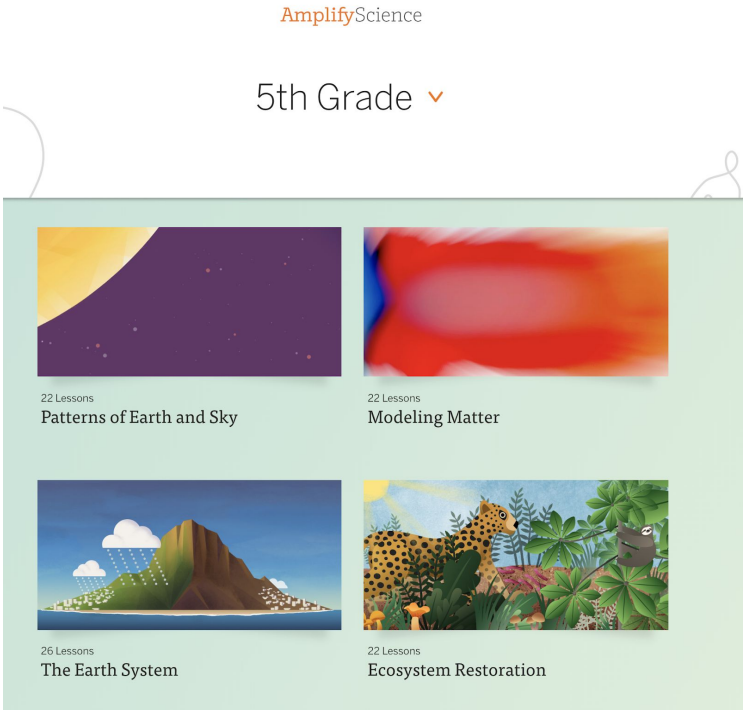
Introduction

This paper describes a study of science outcomes for schools using Amplify Science based on analysis of the Washington Comprehensive Assessment System Grade 5 Science test. The study was designed as a test of the product's theory of action: schools using Amplify Science should keep pace with or even outperform schools using other programs. Specifically, a multiple regression analysis was conducted to investigate the relationship between a school's use of Amplify Science and science performance, while controlling for selection bias using previous scores and school-level demographic variables.

Amplify Science

Amplify Science is a core curriculum for grades K–8 authored by the curriculum experts at UC Berkeley's Lawrence Hall of Science. Teachers using Amplify Science have access to print resources, digital tools, and materials kits for use in the classroom. Amplify Science was designed to keep students interested and engaged while addressing 100% of the Next Generation Science Standards. Four units were specifically designed for 5th grade students: *Patterns of Earth and Sky; Modeling Matter; The Earth System;* and *Ecosystem Restoration*.

Figure 1. Amplify Science 5th Grade units



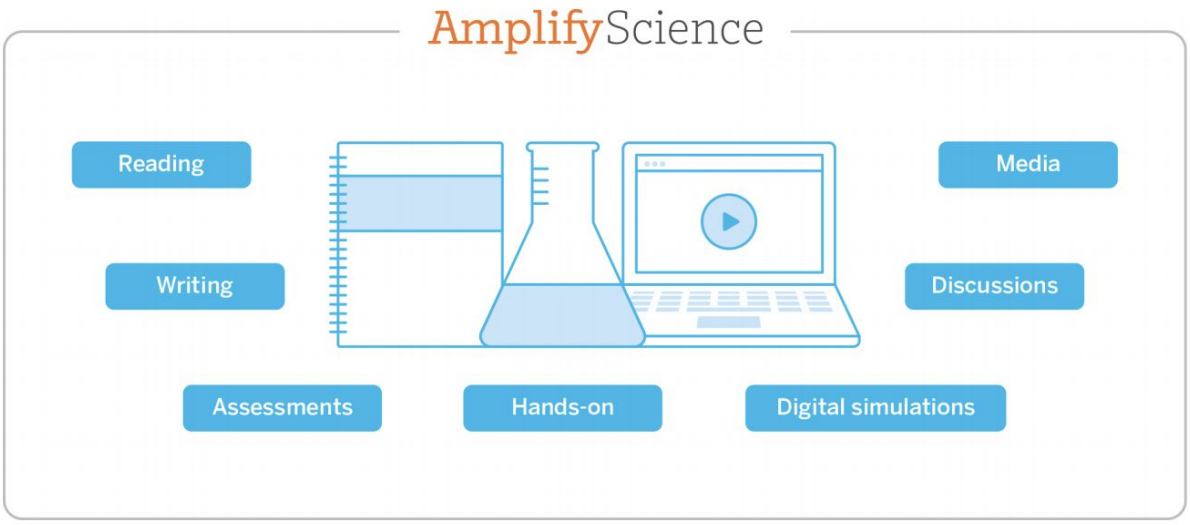
Each unit contains a Teacher's Guide that explains what the unit covers, why these topics are important for students' success in science, and how Amplify Science organizes these topics in a developmental progression. These guides make the logic and pedagogical rationale of each unit available to teachers, in practical language, making it more likely that teachers will implement lessons according to the recommended sequence and be able to track student outcomes in real-time against classroom observations and assessments. For example, this excerpt from the *Patterns of Earth and Sky* Teacher's Guide explains how the models, simulations and readings in the unit facilitate students growth in spatial reasoning ability and the ability to explain phenomena (Lawrence Hall of Science, 2018):

> The spatial reasoning involved in understanding many space science ideas is challenging. In this unit, it involves understanding the position of stars in relation to Earth and the sun, as well as figuring out how Earth's spin and orbit cause us to see different things across a day and across a year. Over the course of the unit, students have repeated opportunities to investigate these patterns, through multiple models. Through the use of these models, students also begin to develop a sense of the large distances and scale of objects in the universe. Access to these ideas through a series of kinesthetic models, physical models, a computer model, and text enables students to have the collection of experiences that are necessary for students to begin to own these ideas. Being able to explain the illustrations on an ancient artifact that shows changes in the sky over time and to figure out what might be on

the missing piece is an intriguing and complex task. It's one that requires knowledge of Earth's daily and yearly movement as well as an understanding of gravity. As such, the problem provides a compelling series of real world phenomena (as captured in the illustrations on the artifact) for students to figure out and explain. In addition, the task of explaining the artifact unifies a set of related and important space science ideas, often taught separately.

Within each unit, teachers access a series of lessons. The lessons provide all of the media, materials and guidance necessary to facilitate teacher-lead and student-to-student discussions, perform digital simulations of processes, set up hands-on investigations with materials shipped in the kits that accompany Amplify Science, perform student assessment, and have students practice reading and writing about science.

Figure 2. The variety of materials contained in the teacher lesson guides



Present study

Amplify Science was designed to improve student outcomes by relying on a strong theory of action for each unit. This study investigates whether use of Amplify Science had a positive impact on student scores. Specifically, this study seeks to answer the following questions:

1. Is use of Amplify Science associated with stronger science performance than use of other programs when controlling for selection bias?
2. If there is a difference in learning outcomes for Amplify Science users, is this difference statistically significant?

Method

5th Grade science outcomes from 2016–2017 were collected from Washington schools that would and would not be using Amplify Science during the 2017–2018 academic year. These outcomes were used as a pretest measure. During the following academic year (2017–2018) teachers in the Amplify Science schools were then given access to the four Amplify Science Units for 5th Grade. 5th grade science outcomes from all of the schools were then collected again in 2017–2018 and used to study the relationship between science performance and use of Amplify Science.

Participants

The treatment group was defined as Elementary Schools that had Amplify Science during the 2017–2018 academic year and had access to all four units for 5th Grade. Schools which also had access to Amplify's middle school science curriculum were excluded from the treatment group because this program allows both teachers *and students* to access the digital platform independently. By excluding users of both Amplify elementary and Amplify middle school curriculum products, the relationship between use of Amplify's elementary school science curriculum and science performance could be investigated separately. Thus for the purpose of this study, an elementary school, was considered any public school that included Grade 5, but did not also include Grades 6, 7, or 8. The grade spans for elementary schools included in this study, in order of frequency, were: K–5, PK–5, 3–5, 1–5, 2–5 and 4–5, with K–5 and PK–5 schools making up 97% of the sample. The control group was defined as elementary schools that did not have Amplify Science during the 2017–2018 academic year.

Table 1. Number of treatment and control units

School type	Number of schools
Elementary schools with Amplify Science	33
Elementary schools without Amplify Science	657

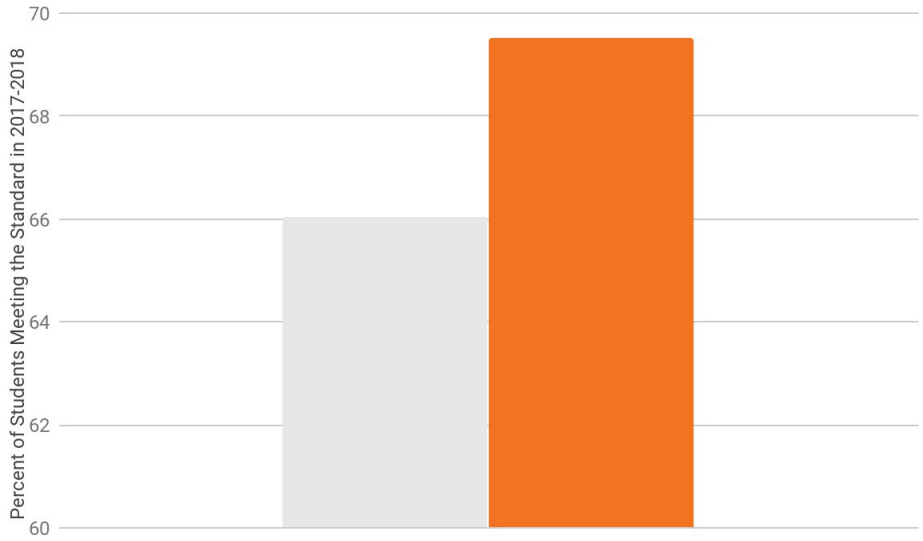
Instruments

Performance on the Washington Comprehensive Assessment System (WCAS) science test for 5th Grade was used as both a pretest and posttest measure. Scores from the 2016–2017 Academic year were used as the pretest and scores from the 2017–2018 academic year, after schools had access to the Amplify Science units for at least 8 months, were used as a posttest. The WCAS is a large-scale standardized assessment developed by the state of Washington to place students into one of four achievement levels that outline which types of academic tasks a student can master and whether they have met the standards adopted by the state. Students who attain Levels 3 and 4 are considered to have met the standard, while students in Levels 1 and 2 have not. More information on the Grade 5 WCAS achievement levels is available through the Washington Office of the Superintendent of Education and Instruction. The percent of students in the school meeting the standard according to the WCAS achievement levels was used as the pretest and posttest metric.

Results

5th Grade science performance for the treatment and control groups was compared to determine its relationship to the use of Amplify Science. To account for selection bias, the model included statistical controls for previous year performance, gender, race/ethnicity, and English Language Learner status. After controlling for these variables, Amplify Science was associated with a statistically significant positive effect. Specifically, Amplify Science is associated with an average +3.47 point difference in the % of students meeting the standards for 5th Grade science, as measured by the Grade 5 WCAS Science Assessment in 2017–2018.

Figure 3. Comparison of science test results



To describe this difference in terms of standard deviation (SD) units, the posttest outcome, along with all of the continuous predictor variables (Previous Year Performance, Percent Male, Percent White, Percent Transitional Billingual) were standardized and the model was re-estimated. This second set of estimates shows that Amplify Science is associated with an average +0.19 SD difference in the % of students meeting the standards for 5th Grade science. Effect sizes of +0.19 are classified as small (Cohen, 1998) or medium (Kraft, 2018). This effect size can be translated into differences in percentiles as well. This means that for an average school not using Amplify Science whose scores were at the 50th percentile, they may have been able to achieve scores at the 57th percentile with Amplify Science, a change of 7 percentiles.

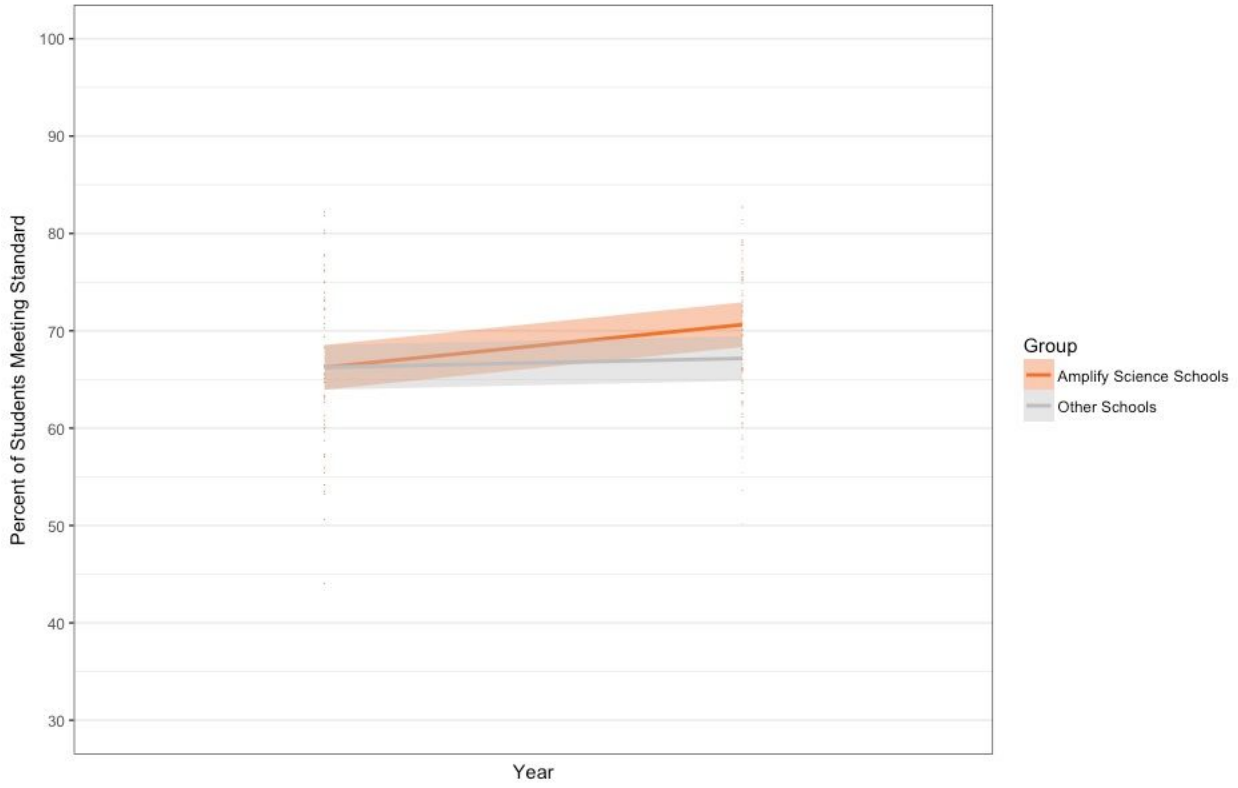
Table 2. Regression model showing how performance is related to Amplify Science

	Coefficient estimates for model with outcome and continuous predictors on original scale	Coefficient estimates for model with standardized outcome and standardized continuous predictors
Amplify Science	3.47*	0.19*
	(1.62)	(0.09)
Previous Year Performance	0.76***	0.79***
	(0.02)	(0.03)
Male	-0.11	-0.02
	(0.14)	(0.02)
White	-0.04	-0.05
	(0.03)	(0.03)
Transitional Billingual	-0.16***	-0.13***
	(0.04)	(0.03)
Y-intercept	16.87*	-0.02*
	(7.52)	(0.02)
Adjusted R2	0.75	0.75
N	690	690

*p<.05, **p<.01, ***p<.001, (standard errors in parentheses)

To help visualize the effect described by these estimates, the expected average growth trajectories for each group have been plotted, starting with an average of 64.70% of students meeting the standard for schools in both groups in Year 1. Figure 4 shows that on average, Amplify schools would have 69.51% of students meeting the standard in year 2, whereas other schools would also show growth, but have 66.04% of students meeting the standard in year 2, a difference of 3.47 percentage points in year 2 outcomes. Figure 4 also illustrates that this difference is significant and enough to distinguish Amplify schools outside of the margin of error.

Figure 4. Model average growth for Amplify Science schools and other schools



Discussion

The purpose of this study was to investigate the following questions:

1. Is use of Amplify Science associated with stronger science performance than use of other programs when controlling for selection bias?
2. If there is a difference in learning outcomes for Amplify Science users, is this difference statistically significant?

The results show that use of Amplify Science is associated with an average 3.47 point gain in the % of students meeting the Grade 5 science standards (effect size +0.19), a result which was statistically significant. Although this design and these results would meet the Tier 3, promising, guidelines for ESSA as a correlational study, using a standardized measure, with controls for selection bias, that shows a statistically significant positive effect, ESSA does not currently rate science programs.

The What Works Clearinghouse (WWC), which does currently rate science programs, requires that studies be either Randomized Control Trials, Quasi-experiments or Regression Discontinuity Designs in which the treatment is

assigned based on previous outcomes (i.e. students with low reading scores being placed in an intervention assigned to help struggling readers). Quasi-experiments, in which the treatment and control group are matched, but are not randomly assigned and are also not assigned based on previous outcomes, are eligible to meet the WWC group design standards with reservations. However, had the 33 treatment schools been matched with control schools based on the 2016–2017 outcomes and demographics, the resulting number of analysis units would have been 66. A power analysis would reveal that with a sample size of 66, we would be unlikely to detect effect sizes of the magnitude typical for educational interventions. Over the next few years, growth in the number of schools using Amplify Science is increasing, and attention is being paid to matching treatment schools to controls. The promising evidence from the current study suggests that the results from these larger quasi-experiments might produce stronger positive evidence for the positive impact of Amplify Science.

References

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Institute of Education Sciences (2017). What Works Clearinghouse Standards Version 4.0. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_standards_handbook_v4.pdf

Kraft, M.A. (2018). Interpreting Effect Sizes of Educational Interventions [Brown University Working Paper]. Retrieved from https://scholar.harvard.edu/files/mkraft/files/kraft_2018_interpreting_effect_sizes.pdf

Lawrence Hall of Science (2018). Patterns of Earth and Sky: Analyzing Stars on Ancient Artifacts - Teachers Guide. New York, NY: Amplify Education, Inc.

United States Department of Education (2016). Non-Regulatory Guidance: Using Evidence to Strengthen Education Investments. Retrieved from https://ed.gov/policy/elsec/leg/essa/guidanceuseseinvestment.pdf